

University of Groningen

**Complete structure of the  $\alpha$ B-crystallin gene: Conservation of the exon-intron distribution in the two nonlinked  $\alpha$ -crystallin genes**

Quax-Jeuken, Y.; Quax, W.; Van Rens, G.; Khan, P. Meera; Bloemendal, Hans

*Published in:*

Proceedings of the National Academy of Sciences of the United States of America

*DOI:*

[10.1073/pnas.82.17.5819](https://doi.org/10.1073/pnas.82.17.5819)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

1985

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Quax-Jeuken, Y., Quax, W., Van Rens, G., Khan, P. M., & Bloemendal, H. (1985). Complete structure of the  $\alpha$ B-crystallin gene: Conservation of the exon-intron distribution in the two nonlinked  $\alpha$ -crystallin genes. *Proceedings of the National Academy of Sciences of the United States of America*, 82(17), 5819-5823. <https://doi.org/10.1073/pnas.82.17.5819>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Complete structure of the $\alpha$ B-crystallin gene: Conservation of the exon–intron distribution in the two nonlinked $\alpha$ -crystallin genes

(crystallins/3' conservation/chromosome mapping)

YVONNE QUAX-JEUKEN\*, WIM QUAX\*, GEERT VAN RENS\*, P. MEERA KHAN†, AND HANS BLOEMENDAL\*‡

\*Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen, The Netherlands; and †Department of Human Genetics, Sylvius Laboratories, University of Leiden, Wassenaarseweg 72, 2333 AL Leiden, The Netherlands

Communicated by L. N. M. Duysens, April 25, 1985

**ABSTRACT** We isolated bovine complementary DNA clones for the  $\alpha$ A- and  $\alpha$ B-crystallin subunits. The  $\alpha$ B cDNA clone was used to isolate an  $\alpha$ B-crystallin gene. This gene, derived from hamster, occurs as a single copy in the genome and is 3.2 kilobases long. The coding sequences are spread on three exons with a total length of 709 nucleotides. The exon–intron distribution of the hamster  $\alpha$ B-crystallin gene is similar to that of the  $\alpha$ A-crystallin gene except for the 69 nucleotides that specify the 23 “insert” residues of the  $\alpha$ A<sup>ins</sup> chain by means of differential splicing. The 3' noncoding region of the  $\alpha$ B mRNA (140 bases), which is short compared with the  $\alpha$ A mRNA (520 bases), shows a remarkable homology between calf and hamster. Both  $\alpha$ -crystallin cDNA clones have been used to assign the chromosomal location of the corresponding human genes with the aid of somatic cell hybrids. It is shown that the single-copy  $\alpha$ A- and  $\alpha$ B-crystallin genes are located on different chromosomes.

Among the structural proteins of the mammalian lens,  $\alpha$ -crystallin is a major component with a molecular weight of  $\approx$ 800,000, which is composed of acidic and basic polypeptides. Of these subunits, only two are primary gene products, named  $\alpha$ A2 and  $\alpha$ B2, varying in ratio and in rate of synthesis during differentiation of the lens cell. All other  $\alpha$ -crystallin subunits arise from  $\alpha$ A2 and  $\alpha$ B2 by post-translational modification (1–3). The hamster  $\alpha$ A-crystallin gene, whose structure has been elucidated recently (4), consists of four exons. Exon 2, which encodes the so-called insert peptide, characteristic for some rodents (5, 6), arises by differential splicing (4–7).

To obtain structural evidence for the different expression patterns of the two  $\alpha$ -crystallin genes, we isolated  $\alpha$ A- and  $\alpha$ B-crystallin cDNA clones from a cDNA bank of bovine lens mRNAs (8). Previously,  $\alpha$ A-crystallin cDNA clones had been isolated for a few other species (9–11, §). Quite surprisingly, however, an  $\alpha$ B-specific cDNA clone had never been found in spite of the fact that several research groups pursued studies on molecular cloning of the crystallin genes. We have now also isolated and characterized by sequence analysis the hamster  $\alpha$ B-crystallin gene and performed comparative studies with the hamster  $\alpha$ A-crystallin gene (4). Moreover, we investigated the chromosomal location of the  $\alpha$ -crystallin genes to address the question of whether these genes are linked on the genome, like the  $\delta$ - and  $\gamma$ -crystallin families; such linkage might have implications for the regulation of expression.

## MATERIALS AND METHODS

**Isolation and Identification of Bovine  $\alpha$ A- and  $\alpha$ B-Crystallin cDNA Clones.** Isolation and cloning of bovine lens mRNAs, the screening of the recombinant clones in a hybridization

selection assay, and identification of the translational products by one- and two-dimensional gel electrophoresis was done as described (8). This resulted in the isolation of the  $\alpha$ A- and  $\alpha$ B-crystallin clones pBL $\alpha$ A-1 and pBL $\alpha$ B-1.

**Construction and Screening of a Hamster Genomic Library.** A partial gene library from a Syrian gold hamster was constructed. *Sau*3A restriction fragments ranging from 15 to 20 kilobases (kb) were cloned in Charon 28 as vector (12). Screening was done with the probes pBL $\alpha$ A-1 and pBL $\alpha$ B-1. Hybridization was carried out at 42°C for 16 hr in 50% formamide/5 $\times$  NaCl/Cit/1 $\times$  Denhardt's solution/20 mM sodium phosphate, pH 6.8/5 mM EDTA/single-stranded salmon sperm DNA (100  $\mu$ g/ml) (1 $\times$  NaCl/Cit = 0.15 M NaCl/0.015 M Na citrate; 1 $\times$  Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone). Washings were at 42°C for 3 hr with the hybridization solution at 55°C for 15 min in 2 $\times$  NaCl/Cit/0.1% NaDodSO<sub>4</sub>. Positive plaques were purified, and characterization was done by restriction enzyme mapping using established methods.

**DNA Sequence Analysis.** Two *Eco*RI fragments, derived from clone  $\lambda$ H $\alpha$ B, of 4.6 kb and 4.2 kb containing the part that hybridizes to our cDNA probe, were subcloned into a suitable pBR322 vector. The inserts of these subclones were digested with the enzymes *Sau*3A, *Hae* III, and *Hpa* II, and the resulting fragments were ligated into M13 mp8, mp9, mp10, or mp11 (13). The chain-termination method (14) was used to sequence the M13 inserts. Analysis was on 40-cm 6% sequencing gels. The gel readings were recorded, edited, and compared by the programs of Staden (15). Search for homology and optimal alignment was done by the program DIAGON (16).

**Cell Hybrids and Genetic Analysis.** The methods for production, isolation, and propagation of the interspecific somatic cell hybrids have been reported (17, 18). Previous reports described the PG/Me hybrids (19–21), the a3G hybrid (22), the a3Y hybrids (23), and the E36 series of hybrids (24) included in the present study. The hybrids were investigated for identifying the individual human chromosomes with G-banding, Giemsa-11, and/or Q-banding techniques (25). A list of 37 chromosome-specific enzyme markers used to screen to hybrids, including the references of the methods used, was reported elsewhere (24).

**DNA Preparation and Southern Hybridization Analysis.** Total DNA of the cell lines was isolated as described (26). Each DNA sample (10  $\mu$ g) was digested with appropriate restriction enzymes, electrophoresed on 0.8% agarose gels, blotted onto nitrocellulose, and hybridized under the condi-

Abbreviations: kb, kilobase(s); bp, base pair(s).

‡To whom reprint requests should be addressed.

§Yasuda, K., Kondoh, H., Okazaki, K., Shimura, Y. & Okada, T. S., Fifth International Congress of Eye Research, October 3–8, 1982, Veldhoven, The Netherlands, p. 38 (abstr.).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

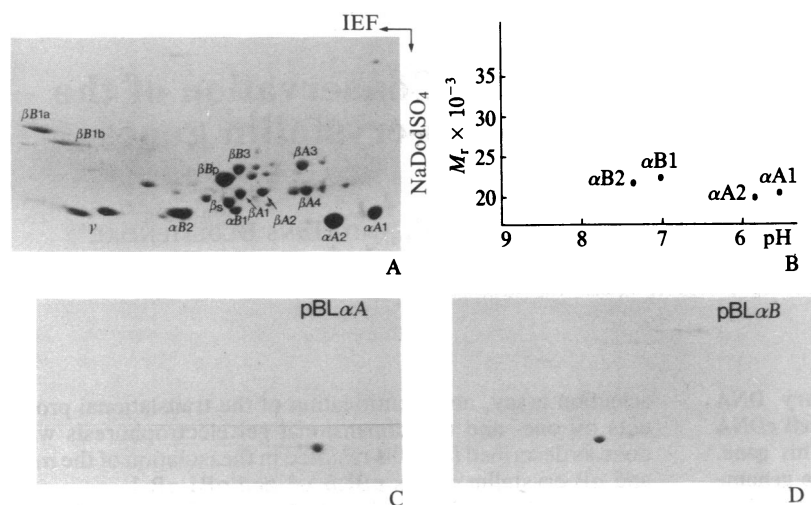


FIG. 1. Identification of clones containing  $\alpha$ -crystallin cDNA sequences by hybridization selection and two-dimensional NaDodSO<sub>4</sub>/polyacrylamide gel electrophoresis. (A) Coomassie blue-stained gel showing the crystallins from calf lens cortex. (B) Schematic representation of  $\alpha$ -crystallin polypeptides of bovine lens.  $\alpha$ A1 and  $\alpha$ B1 are post-translationally modified products of  $\alpha$ A2 and  $\alpha$ B2. (C and D) Fluorographs of gels showing the *in vitro* translation products of mRNAs hybridized to the specific plasmids pBL $\alpha$ A-1 and pBL $\alpha$ B-1, respectively. Identification was accomplished by matching the fluorograph with the cold pattern of the  $\alpha$ -crystallins run on the same gel. IEF, isoelectric focusing.

tions described above. Plasmid inserts of pBL $\alpha$ A-1 and pBL $\alpha$ B-1 were <sup>32</sup>P-labeled by nick-translation.

## RESULTS

**Identification of  $\alpha$ A- and  $\alpha$ B-Crystallin cDNA Clones.** A bovine cDNA library constructed in pBR322 was screened for  $\alpha$ A- and  $\alpha$ B-crystallin cDNAs by a hybridization selection assay (8). The polypeptides synthesized *in vitro* were analyzed by one-dimensional (data not shown) and two-dimensional NaDodSO<sub>4</sub>/polyacrylamide gel electrophoresis and fluorography. Fig. 1 shows the identification of an  $\alpha$ A- and  $\alpha$ B-crystallin cDNA clone by two-dimensional gel coelectrophoresis with the soluble protein fraction of bovine lens. Fig. 1A presents one of these gels stained with Coomassie blue showing the migration of the bovine crystallins used as markers. Fig. 1C and D show the products synthesized *in vitro* from the mRNAs selected by the clones pBL $\alpha$ A-1 and pBL $\alpha$ B-1. The inserts of these  $\alpha$ -crystallin clones were 1000 base pairs (bp) for  $\alpha$ A and 700 bp for  $\alpha$ B. Of the clones tested, 15% contained  $\alpha$ -crystallin specific cDNA ( $\alpha$ A and  $\alpha$ B clones were found at a ratio of 4:1).

**Isolation of the Hamster  $\alpha$ B-Crystallin Gene.** A hamster genomic library constructed in the  $\lambda$  phage Charon 28 (27) was screened for the  $\alpha$ B-crystallin gene by using the nick-translated cDNA clone pBL $\alpha$ B-1 as hybridization probe. After screening  $5 \times 10^5$  recombinants, one positive clone was identified. For restriction enzyme and blotting analysis, two *Pst* I fragments (150 and 500 bp long) were isolated from pBL $\alpha$ B-1 and used as probes. It was found that the coding sequences of the  $\alpha$ B-crystallin gene were located within two *Eco*RI fragments and that the gene is  $\approx 3.2$  kb long flanked by 5.8 kb 5' upstream and 4.6 kb 3' downstream sequences (Fig. 2). To allow a more detailed study of the gene, we made subclones of the two *Eco*RI fragments (4.6 and 4.2 kb) in pBR322, which were used for subsequent sequence analysis.

Southern blot analysis of hamster genomic DNA with the insert of pBL $\alpha$ B-1 as probe revealed the same hybridizing

bands as the blotting analysis of the two *Eco*RI subclones (data not shown). These findings clearly indicate that all hamster genomic  $\alpha$ B-crystallin sequences are located within these two *Eco*RI fragments and that  $\alpha$ B-crystallin is encoded by a single-copy gene.

**Nucleotide Sequence of the  $\alpha$ B-Crystallin Gene.** To define the exon/intron organization of the gene and to examine the structural basis of  $\alpha$ B-crystallin expression, we determined the sequence of the coding parts and flanking sequences (Fig. 3). A comparison of the gene sequences with the reported bovine and human  $\alpha$ B-crystallin amino acid sequences (28, 29) allowed us to locate the different exons. As shown in Figs. 2 and 3, the hamster  $\alpha$ B gene consists of 3 exons and 2 introns. The first exon contains the coding information for amino acids 1–67 as well as 41 bp of the 5' noncoding region. The second exon has a length of 123 bp and encodes residues 68–108 of the  $\alpha$ B chain. The third and last exon is 340 bp long and contains the coding sequence for amino acids 109–175 and the information for the 140 bp of the 3' noncoding region. The two introns of the  $\alpha$ B-crystallin gene are 0.9 kb and 1.6 kb long. The translated exon sequences are almost identical with the known amino acid sequence of the  $\alpha$ B2 subunit of calf and man (28, 29); the few substitutions are indicated in Fig. 3. Residues 39–41 apparently have a high mutation rate because they differ in all species examined.

In accordance with the general GT–AG border nucleotide rule, both introns begin with the dinucleotide GT and end with AG (30). More extensive consensus sequences have been proposed for the splice sites of eukaryotic genes that are transcribed by RNA polymerase II (31). The donor ( $\frac{C}{A}-\frac{G}{G}-\frac{T}{G}-\frac{A}{A}-\frac{T}{G}-\frac{T}{T}$ ) and especially the acceptor [ $\frac{C}{T}$ ]<sub>11</sub>–N– $\frac{C}{T}$ –A–G/G] consensus sequence are reasonably well matched by the exon–intron junctions of the  $\alpha$ B-crystallin gene.

These exon–intron borders were compared with the corresponding regions of the hamster  $\alpha$ A-crystallin gene (4). Since the products of the  $\alpha$ A- and  $\alpha$ B-crystallin genes  $\alpha$ A2

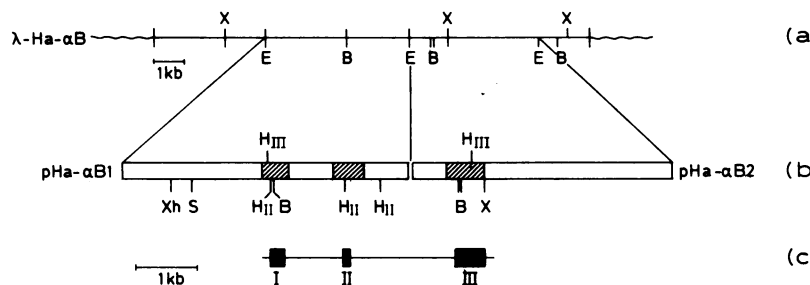


FIG. 2. Physical map of the  $\alpha$ B-crystallin gene. (a) The sites for some restriction enzymes in the  $\lambda$ Ha- $\alpha$ B insert are shown. B, *Bam*HI; E, *Eco*RI; HIII, *Hind*III; HII, *Hinc*II; Xh, *Xho* I; X, *Xba* I; S, *Sac* I. (b) Shaded bars are the regions whose sequence is presented in Fig. 3. (c) Black bars represent the position and size of  $\alpha$ B exons, which are numbered with roman numerals. Note: There is an additional *Sac* I site 0.2 kb upstream from the *Xba* I site.

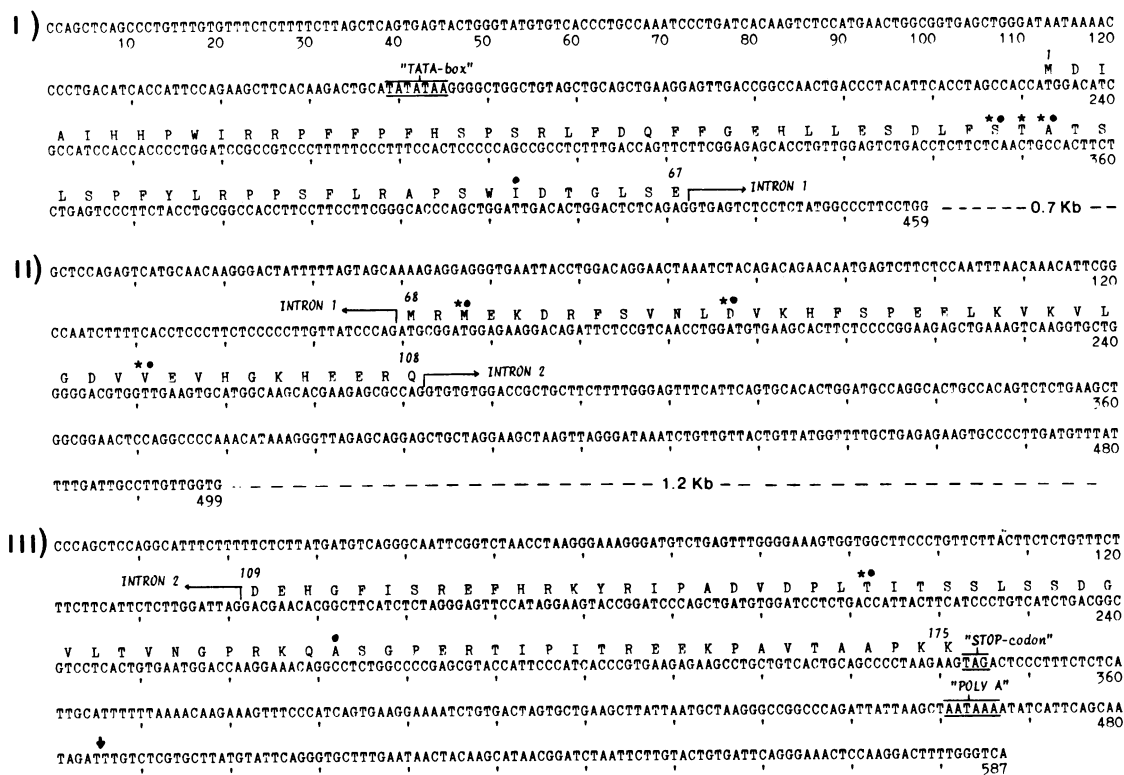


FIG. 3. Sequence of the hamster  $\alpha$ B-crystallin gene. The sequences of all coding regions and flanking nucleotides are shown. Introns are marked by arrows and the amino acid sequence encoded by the exons is given in the single-letter code. Lengths of intron fragments that are not sequenced are indicated. TATA box points to the transcriptional control signal. Arrow at position 485 indicates the poly(A) addition site, which is found by comparison with the bovine  $\alpha$ B-crystallin cDNA clone sequence (Fig. 5). Asterisks and dots mark amino acid substitutions between hamster  $\alpha$ B-crystallin and the known amino acid sequences of calf (28) and human (29), respectively.

and  $\alpha$ B2 have 56% sequence homology (28), they are supposed to be derived from a common ancestral gene by a duplication event, which is thought to have taken place >500 million years ago. As can be seen in Fig. 4, all exon-intron junctions are exactly located at the same sites in both genes, thereby yielding further evidence for the evolutionary relationship of these two genes.

**Flanking Regions.** The 5' and 3' flanking regions were also sequenced because these regions may be important in the regulation of gene expression. At positions 159–165 in Fig. 3, there is a sequence "TATATAA" that obeys the consensus promoter sequence for eukaryotic genes (Goldberg-Hogness box). The transcription initiation site or the cap site could tentatively be assigned to the adenine residue at position 189,

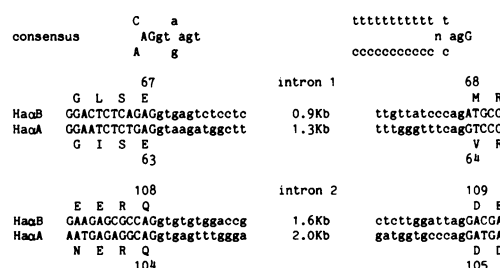


FIG. 4. Exon-intron junctions of the  $\alpha$ B-crystallin gene compared with the corresponding junctions of the hamster  $\alpha$ A-crystallin gene (4). Nucleotides around the splice sites are compared to the consensus donor (Left) and acceptor (Right) sequence (30). Upper and lower lines represent hamster  $\alpha$ B- and hamster  $\alpha$ A-crystallin sequences, respectively. Exon sequences are in capital letters; amino acids bordering the introns are given above and below the corresponding codons. The insert sequence of the  $\alpha$ A<sup>Ins</sup> is not included. The "ag" and "gt" intron consensus sequences are always present, the other nucleotides are slightly different from the consensus.

which is not preceded by the consensus cytosine, but rather by guanine. As for the  $\alpha$ A-crystallin gene of hamster (4) and mouse (7) and the chicken  $\delta$ 2-crystallin gene (32), no consensus CCAAT-box could be detected. The 3' end of the hamster  $\alpha$ B-crystallin gene was determined by comparing the nucleotide sequence of the 3' untranslated region of the gene with the sequence of the bovine cDNA clone pBL $\alpha$ B-1, which contains the complete 3' noncoding region. The polyadenylation signal AATAAA is located 23 nucleotides upstream from the poly(A) addition site. Comparison of the 3' noncoding regions of the bovine cDNA clone and the hamster  $\alpha$ B-crystallin gene clone (Fig. 5) revealed that these regions are equal in length and have >90% sequence homology.

**Mapping of the  $\alpha$ A- and  $\alpha$ B-Crystallin Genes on Human Chromosomes.** The linkage relationship between the two  $\alpha$ -crystallin genes in the genome was studied by investigating the chromosomal location of the  $\alpha$ A- and  $\alpha$ B-crystallin genes by using somatic cell hybrids (rodent-human). We used a panel of five mouse-human hybrids and a panel of nine

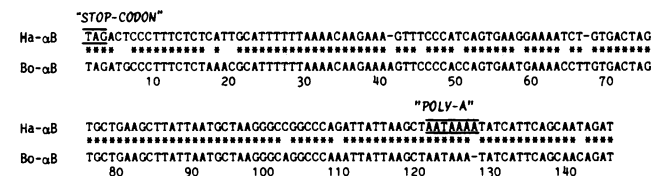


FIG. 5. Comparison of 3' nontranslated regions of the hamster  $\alpha$ B-crystallin gene and bovine  $\alpha$ B cDNA clone pBL $\alpha$ B. The 3' noncoding sequences of hamster and calf are aligned. Common residues are indicated by an asterisk. From the stop codon up to the poly(A) addition site, the length of the 3' noncoding sequence is identical and the homology is >90%.

Table 1. Distribution of human chromosomes in cell hybrids

		Human chromosome																									
Lane no. (Fig. 6)	Clone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	$\alpha$ A	$\alpha$ B	
H	a3Y 14-2	+		+	+	+			+		+	+	+	+	+	+	+			+	+	+	+		+	+	
I	a3Y 13-1	+		+				+	+			+	+	+	+	+	+	+		+	+	+		+	+	+	
J	a3G 5-2	+				+	+							+		+	+	+	+		+		+			+	
K	E36 33.7-2	+	+	+	+		+		+		+				+		+	+	+	+				+		+	
L	E36 33.11-2	+	+					+					+	+	+	+	+	+			+	+		+	+	+	
M	E36 10 CB 1-2	+	+	+	+			+	+			+	+	+			+	+				+		+	+	+	
N	E36 10 CB 22B2				+				+	+	+	+					+					+		+	+	+	
P	E36 78-3-13-2	+	+			+						+				+	+		+	+		+		+	+	+	
Discordance $\alpha$ A		3	4	4	4	5	8	3	4	5	5	1	2	3	4	3	2	4	7	3	4	0	6	2			
$\alpha$ B		1	4	4	4	5	6	5	2	7	5	3	4	3	4	3	0	2	5	5	4	2	6	2			

Cell lines H-P are Chinese hamster-human somatic hybrids. +, Presence of human chromosome containing  $\alpha$ A and  $\alpha$ B sequences. The lowest discordance is with chromosome 21 for  $\alpha$ A and with chromosome 16 for  $\alpha$ B. Cell line O is omitted from this table, because the precise chromosome content of this line could not be identified.

Chinese hamster-human hybrids. The cryopreserved cell lines, which were screened for the human chromosomes and chromosome-specific markers during previous studies (19, 22-24), were brought back to culture, and the cell pellets of each hybrid line were divided into two portions. One part was used to verify the human chromosome content by retesting for chromosome-specific enzyme markers (Table 1). The other part was used to isolate total DNA, which had been digested with an appropriate restriction enzyme and, after

electrophoresis and blotting, was hybridized to the nick-translated insert of pBL $\alpha$ A-1 and pBL $\alpha$ B-1. For the  $\alpha$ A-crystallin probe, it was found that in the case of hamster-human cell hybrids, the restriction enzyme *Bam*HI gave a good resolution of the hamster and human bands. The 4.0-kb human *Bam*HI fragment is found in lanes H, I, L, M, N, and P (Fig. 6 *Upper*). Comparison of these results with the chromosome information of the hybrids (Table 1) indicated that the  $\alpha$ A-crystallin gene is in chromosome 21. The use of a panel of five mouse-human cell lines also was in concert with the assignment of the  $\alpha$ A-crystallin gene to chromosome 21 (data not shown). The  $\alpha$ B probe gave a good resolution between the human and hamster bands with *Sac*I. The human 10-kb *Sac*I fragment is found in lanes H-P (Fig. 6 *Lower*). Comparison of these results with Table 1 and calculation of a discordance score indicates that the  $\alpha$ B-crystallin gene is presumably in chromosome 16. Some doubt remains about the hybridization signal in the cell line in lane L. Only a very faint band with a different migration behavior could be detected after hybridizing with the  $\alpha$ B-crystallin probe. Screening of another panel of cell lines (mouse-human) did not give conclusive results. From these results, however, it is clear that both  $\alpha$ -crystallin genes segregate with different chromosomes, excluding the possibility of a close physical linkage of these genes on the same chromosome.

## DISCUSSION

The highly conserved slowly evolving crystallins have demonstrated their usefulness for the study of molecular evolution (34). The  $\alpha$ -crystallins have the lowest mutation rate among crystallins, only 3 amino acid replacements per 100 residues in 100 million years.

At the DNA level, the  $\alpha$ B-crystallin is also highly conserved, surprisingly not only for the protein-encoding nucleotides. Alignment of the exon sequences of the hamster  $\alpha$ B-crystallin gene with the sequence of the bovine  $\alpha$ B-crystallin cDNA revealed a high degree of homology in the 3' noncoding region (Fig. 5)—only 3 single base deletions and 15 substitutions in a sequence of 140 bp. Assuming an equal rate of mutation after divergence and a neutral mutation rate for 3' noncoding regions, which was estimated to be 0.7% per million years (35), this degree of difference would imply that divergence took place  $\approx 0.7 \times 10^7$  years ago. However, the last common ancestor of hamster and calf existed probably  $6 \times 10^7$  years ago. Therefore, considerable selective constraint must have maintained this sequence homology between the 3' untranslated regions after species divergence. On the contrary, the unusually long 3' noncoding part of the  $\alpha$ A-

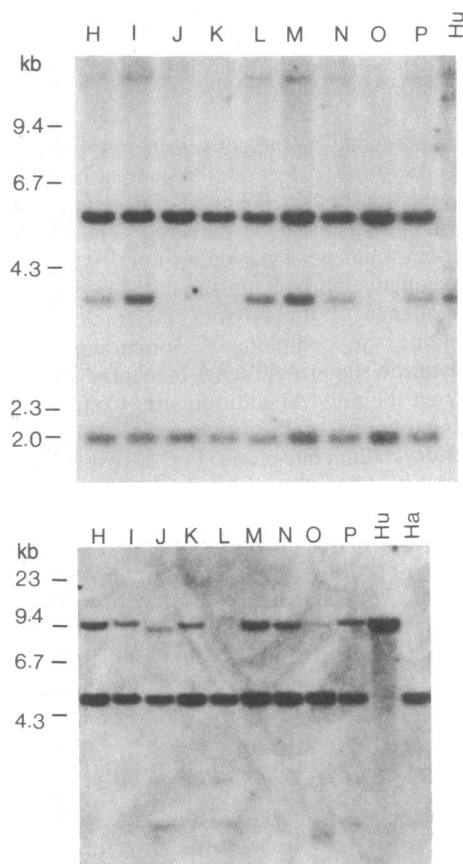


FIG. 6. Assignment of the  $\alpha$ A- and  $\alpha$ B-crystallin genes. DNAs isolated from somatic cell hybrids between hamster and man were digested with *Bam*HI and hybridized with the insert of pBL $\alpha$ A (*Upper*) or digested with *Sac*I and hybridized with the insert of pBL $\alpha$ B (*Lower*). The chromosome content of each cell line (lanes H-N, P) is described in Table 1. Control DNAs: Hu, human; Ha, hamster.

crystallin mRNA (520 bases for the hamster) does not show this evolutionary sequence conservation (4), since there is already 20% difference in this region comparing the closely related species hamster and mouse. Homologous 3' untranslated regions have been reported for members of several multigene families encoding cytoskeletal proteins, actins (36), tubulins (37), and also for the intermediate filament protein vimentin (38). It has been suggested that these conserved 3' noncoding regions might be necessary for the secondary structure of the RNA or for protein-nucleic acid interactions.

Comparative studies of the hamster  $\alpha$ A- and  $\alpha$ B-crystallin genes revealed almost no homology in the noncoding parts. Only consensus sequences like the "TATA" box and the poly(A) addition signal are conserved in addition to several short stretches of homology. A consensus "CCAAT" box could be detected neither in the  $\alpha$ A-crystallin gene nor in the  $\alpha$ B-crystallin gene.

Initial experiments revealed a potentially important DNA sequence for the expression of the murine  $\alpha$ A-crystallin gene between -85 and -400 bp from the cap site (39). The sequence of the  $\alpha$ B-crystallin gene is determined up to -140 bp from the TATA box. Therefore, it might be possible that stretches of sequences that are involved in the regulation of transcription and in the tissue-specific expression of the  $\alpha$ -crystallin genes are situated upstream from the sequence, as depicted in Fig. 3. Further experiments with a suitable expression system are needed to study the promoter regions of the  $\alpha$ -crystallin genes in more detail. Another interesting feature of the  $\alpha$ -crystallins is their homology with the small heat shock proteins (hsp) of *Drosophila* (40), *Caenorhabditis*, and soybean (41, 42). Alignment of the protein sequences of hamster  $\alpha$ B-crystallin and *Drosophila* hsp26 showed that the homologous parts of these proteins are restricted to the second and third exon of the  $\alpha$ B-crystallin gene (data not shown). The homology starts at residue 68 of the  $\alpha$ B chain, the 5' end of the second exon. These confined stretches of homology were also found for the hamster  $\alpha$ A-crystallin gene (4). It may be suggested that the greatest part of the  $\alpha$ -crystallin genes has originated by duplication from ancestral heat shock genes. Combination of different exons might have taken place after this duplication. These suppositions would be in agreement with the hypothesis that new proteins may arise by combining exons encoding separate structural and functional domains, as proposed by Gilbert (43). The possible function of the heat shock protein-like domain in eye lens proteins is an intriguing object of future studies.

Another aim of our studies was to address the question of how the expression of the  $\alpha$ -crystallin genes is regulated on the genome. Since  $\alpha$ A- and  $\alpha$ B-crystallin together form higher aggregates, it might be supposed that their expression is regulated coordinately. Linkage of the  $\alpha$ -crystallin genes on the genome might be essential in that respect. Several crystallin families are linked on the genome. The  $\delta$ -crystallin genes are only 4.2 kb apart on the chicken genome (44), and the  $\gamma$ -crystallin genes are linked on the human chromosome 2 (33). Our results clearly show that the  $\alpha$ -crystallin genes are not located on the same chromosome, although additional experiments are needed to verify the assignment of the  $\alpha$ B-crystallin gene. The  $\alpha$ A-crystallin gene is assigned to chromosome 21. In this respect, the observed trisomy of chromosome 21 in patients with Down syndrome, who often suffer from juvenile cataract, is striking and may point to a correlation between cataract and abnormal crystallin expression.

We thank Dr. W. de Jong for his comments and Dr. G. Berbers for the two-dimensional gel electrophoresis. The present investigations

have been carried out in part under the auspices of The Netherlands Foundation for Chemical Research (SON) and with financial aid of The Netherlands Organization for the Advancement of Pure Research (ZWO).

- Harding, J. J. & Crabbe, M. J. C. (1984) in *The Eye*, ed. Davson, H. (Academic, New York), 3rd Ed., pp. 207-492.
- Bloemendal, H. (1977) *Science* **197**, 127-138.
- Bloemendal, H. (1982) *Crit. Rev. Biochem.* **12**, 1-39.
- van den Heuvel, R., Hendriks, W., Quax, W. & Bloemendal, H. (1985) *J. Mol. Biol.*, in press.
- Cohen, L. H., Westerhuis, L. W., De Jong, W. W. & Bloemendal, H. (1978) *Eur. J. Biochem.* **89**, 259-266.
- de Jong, W. W., Cohen, L. H., Leunissen, J. A. M. & Zweers, A. (1980) *Biochem. Biophys. Res. Commun.* **96**, 648-655.
- King, C. R. & Piatigorsky, J. (1983) *Cell* **32**, 707-712.
- Quax-Jeuken, Y., Janssen, C., Quax, W., van den Heuvel, R. & Bloemendal, H. (1984) *J. Mol. Biol.* **180**, 457-472.
- Dodermont, H. J., Andreoli, P. M., Moormann, R. J. M., Ramaekers, F. C. S., Schoenmakers, J. G. G. & Bloemendal, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5320-5324.
- King, C. R., Shinohara, T. & Piatigorsky, J. (1982) *Science* **215**, 985-987.
- Tomarev, S. I., Zinovieva, R. D., Dolgilevich, S. M., Krayev, A. S., Skryabin, K. G. & Gause, G. G. (1983) *FEBS Lett.* **162**, 47-51.
- Rimm, D. L., Horness, D., Kucera, J. & Blattner, F. R. (1980) *Gene* **12**, 301-309.
- Messing, J., Crea, R. & Seeburg, P. (1981) *Nucleic Acids Res.* **9**, 309-321.
- Sanger, F., Coulson, A. R., Barrel, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161-178.
- Staden, R. (1979) *Nucleic Acids Res.* **6**, 2601-2610.
- Staden, R. (1982) *Nucleic Acids Res.* **10**, 2951-2961.
- Westerveld, A., Visser, R. P. L. S., Meera Khan, P. & Bootsma, D. (1971) *Nature (London)* **234**, 20-24.
- Pontecorvo, G. (1975) *Somatic Cell Genet.* **1**, 391-395.
- Meera Khan, P., Wijnen, L. M. M., Wijnen, J. Th. & Grzeschik, K. H. (1983) *Biochem. Genet.* **21**, 123-133.
- Wijnen, L. M. M., Grzeschik, K. H., Pearson, P. L. & Meera Khan, P. (1977) *Hum. Genet.* **37**, 271-278.
- Wijnen, L. M. M., Grzeschik, K. H., Pearson, P. L. & Meera Khan, P. (1977) *Hum. Genet.* **39**, 339-344.
- Meera Khan, P., Wijnen, L. M. M., Hagemeijer, A. & Pearson, P. L. (1984) *Cytogenet. Cell Genet.* **38**, 112-115.
- Herbschleb-Voigt, E., Pearson, P. L., Vossen, J. M. & Meera Khan, P. (1981) *Hum. Genet.* **56**, 379-386.
- Herbschleb-Voigt, E., Grzeschik, K. H., Pearson, P. L. & Meera Khan, P. (1981) *Hum. Genet.* **59**, 317-323.
- Pearson, P. L. & van Egmond-Cowan, A. M. M. (1976) in *New Techniques in Biophysics and Cell Biology*, eds. Pain, R. H. & Smith, B. J. (Wiley, London), pp. 213-239.
- van der Putten, H., Terwindt, E., Berns, A. & Jaenisch, R. (1979) *Cell* **18**, 109-116.
- Quax, W., van den Broek, L., Vree Egberts, W., Ramaekers, F. & Bloemendal, H. (1985) *Cell*, in press.
- van der Ouderaa, F. J., de Jong, W. W., Hilderink, A. & Bloemendal, H. (1974) *Eur. J. Biochem.* **49**, 157-168.
- Kramps, J. A., de Man, B. M. & de Jong, W. W. (1977) *FEBS Lett.* **74**, 82-84.
- Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459-472.
- Sharp, P. A. (1981) *Cell* **23**, 643-646.
- Piatigorsky, J. (1984) *Mol. Cell. Biochem.* **59**, 33-56.
- den Dunnen, J. T., Jongbloed, R. J. E., Geurts van Kessel, A. H. M. & Schoenmakers, J. G. G. (1985) *Hum. Genet.*, in press.
- de Jong, W. W. (1981) in *Molecular and Cellular Biology of the Lens*, ed. Bloemendal, H. (Wiley, New York), pp. 221-278.
- Perler, R., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. & Dodgson, J. (1980) *Cell* **20**, 555-566.
- Ponte, Ph., Gunning, P., Blau, H. & Kedes, L. (1983) *Mol. Cell. Biol.* **3**, 1783-1791.
- Cowan, N. J., Dobner, P. R., Fuchs, E. V. & Cleveland, D. W. (1983) *Mol. Cell. Biol.* **3**, 1738-1745.
- Quax, W., Vree Egberts, W., Hendriks, W., Quax-Jeuken, Y. & Bloemendal, H. (1983) *Cell* **35**, 215-223.
- Chepelinsky, A. B., King, C. R., Zelenka, P. S. & Piatigorsky, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2334-2338.
- Ingolia, T. D. & Craig, E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2360-2364.
- Russnak, R. H., Jones, D. & Peter, E. (1983) *Nucleic Acids Res.* **11**, 3187-3205.
- Schoffl, F., Rascke, E. & Nagao, R. T. (1984) *EMBO J.* **1**, 2491-2497.
- Gilbert, W. (1978) *Nature (London)* **271**, 501.
- Hawkins, J. W., Nickerson, J. M., Sullivan, M. A. & Piatigorsky, J. (1984) *J. Biol. Chem.* **259**, 9821-9825.